DOCUMENT RESUME

ED 130 647 IR 004 186

AUTHOR	Gekoski, William L.
TITLE	An Investigation of One-week Reliability of the Michigan Restricted Association Norms. Analysis of Meaning with Restricted Association Tests Report Number 5.
INSTITUTION	Michigan Univ., Ann Arbor. Dept. of Psychology.
SPONS AGENCY	National Inst. of Mental Health (DHEW), Bethesda, Md.
PUB DATE	20 Aug 65
GRANT	NIMH-MH-07619-01AI
NOTE	19p.; Archival document
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS	*Association Tests; College Students; Higher Education; *Reliability; *Test Reliability
IDENTIFIERS	*Michigan Restricted Association Norms

ABSTRACT

To test the reliability of four tasks of the Michigan Restricted Association Norms--Superordinates, Parts, Following Words, and Free Associations--30 subjects were asked to respond to 33 stimulus items under each of the four instructions and to perform identical tasks 7-14 days later. Comparisons of the intra-individual reliability and inter-individual response distributions were made between tasks, frequency of stimuli, and sessions. Distributions differed between tasks but not between stimulus frequencies or sessions. Reliability was highest for Superordinates, slightly lower for Free Associations and Parts, and lowest for Following Words. The absolute magnitude of reliability was regarded as sufficiently high though subject to purpose-specific evaluation. Attempts also were made to compare the differences in reliability with differences in the response distribution of the norms. For Restricted but not for Free Associations the strength of the primary responses in the Michigan Norms correlates with the reliability coefficients. (KB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED130647

The University of Michigan

Department of Psychology

An Investigation of One-week Reliability of the Michigan Restricted Association Norms

by

William L. Gekoski

Analysis of Meaning with Restricted Association Tests

Report Number 5

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

20 August 1965

Supported by the National Institute of Mental Health

Grant Number MH 07619 - OIAI

Principle Investigator: Klaus F. Riegel

IR 004 186

An Investigation of One-week Reliability of the Michigan Restricted Association Norms

William L. Gekoski

Early work with free word association was instigated by a desire to establish norms which could prove useful in clinical diagnosis. This was the specified intent of the efforts by Jung (1918), Kent-Rosanoff (1910), Rosanoff and Rosanoff (1913), and Woodrow and Lowell (1916). For application in the clinical setting, however, the norms were unsuccessful as the distinctions based upon them were so gross that nothing more was contributed to diagnoses made through the use of observational-interview methods and/or through other testing devices.

The emphasis has shifted away from the clinical use of free association norms, in favor of their use in experimental situations in which associative and/or mediational processes are being investigated. Jenkins and Russell (1960), Palermo and Jenkins (1965), and Riegel (1965) have reviewed and compared the eight sets of free association norms for adults and children collected since 1910. Thus, for controlling association strength in paired associate learning, for selecting stimuli for mediation or concept identification experiments, for inquiries into semantic structure, and for research concerned with the interaction of language and various levels of cognitive processes, free association norms have proven increasingly useful.

Because traditional estimates of reliability and validity are difficult or even impossible to ascertain, the adequacy of the existing free association norms has never been seriously analyzed. The reliability of the norms has rested almost entirely upon the large sample size (500 to 1000 Ss). Inferences concerning validity have been drawn from the high degree of similarity found to exist between the norms collected at different times over a fifty year span of time (Gekoski, 1964).

By comparison with the free word associations the amount of research dealing with restricted, or controlled, associations has been negligible. The earliest data were

obtained from relatively small samples (Watts, 1905; Wreschner, 1907; May, 1917) and were intended merely to aid in the selection of stimuli for various experimental purposes, e.g., investigations of differential response latencies as a function of task instructions. Recently restricted association norms have been compiled (DiVesta, 1964; Jenkins and Palermo, 1964a,b; Martin and Clifton, 1964), however, each of these studies has been concerned with a few specific types of restrictions only, and emphasizes certain aspects about these types of restrictions, in an effort to account for the nature of free associations. The first systematic, normative study of restricted association, The Michigan Restricted Association Norms, were collected by Riegel (1965). These norms were obtained from the responses of one hundred college students to each of two hundred stimulus words under each of sixteen different restricted task instructions.

The present study has been undertaken in order to investigate and quantitatively estimate the reliability of four tasks of The Michigan Restricted Association Norms (Superordinates, Parts, Following Words, and Free Associations) over a relatively short period of time (7 to 14 days). Furthermore, the dependency of the reliability on the word frequency of the stimulus and the type of restrictions has been analyzed.

The concept of reliability raises some particular problems in the case of word associations and thus may be construed in at least two different ways. First, the norms might be said to be reliable to the degree to which different random samples yield similar response distributions. Second, the norms might be said to be reliable to the extent to which two response distributions obtained from the same Ss at different times are similar. It is in this latter sense that reliability is analyzed in this study. A forthcoming investigation will deal with reliability in the former sense. Reliability studies on longer time intervals are in progress and when completed will be compared to the data presented here.

The present study analyzes the degree to which each S gives the same responses to the same stimuli on the two sessions (intra-individual reliability). Also, if the Norms are reliable it would be expected, in the cases where a S does not give the same response on both sessions that these non-identical responses would be high, or at least present, in the distribution of responses obtained from all Ss. Thus the degree to which these distributions are alike for the two sessions (inter-individual reliability) is investigated.

The single experimental hypothesis of this study concerns differences in reliability as a function of the tasks. Superordinates are expected to have the greatest reliability, Following Words responses the least, the the reliability of Parts is expected to fall in between. This hypothesis was formulated on the assumption that reliability would be a function of the size of the class of possible responses available in the language. No statement was made regarding the reliability of Free Association relative to the restricted tasks. In addition to the question of differential reliability as a function of task, the absolute degree of reliability under each task instruction is of importance for it will allow inferences as to the suitability of employing the norms for various purposes as well as for defining a reference point to which the overlap of any two groups may be compared.

Method

Ss were twenty-two females and eight males enrolled in an undergraduate course at The University of Michigan. All Ss were between twenty and twenty-five years old.

Thirty-three stimuli were selected from The Michigan Restricted Association Norms (Riegel, 1965) and include 21 nouns, six verbs and six adjectives. Eighteen words were of high (AA-words) and 15 of low frequency (<AA-words) according to the general count of Thorndike-Lorge (1944). The 33 stimuli employed in this study are as follows:

AFRAID, CLOSED, COMFORT, COMMAND, DEVOUR, EAGLE, FLAT, FLY, GLASS, HAND, INSECT, JOY, KNIFE, LIGHT, MANSION, METROPOLIS, MINOR, MOUNTAIN MOUTH, PATH, PEOPLE, PREACH, RELIGION, SICKNESS, SIT, SMALL, SPARROW, STREET, TABLE, TOBACCO, TROUBLE, WOMAN, and WOOLLY.

In the task of free word associations, Ss were asked to respond with the first word that came to their minds after reading the stimulus. The restricted association categories employed were Superordinates (a logical task), Parts (an infralogical, or physical, task), and Following Words (a grammatical task). Below are the verbatim instructions to Ss:

> The following are tasks of restricted associations. You have received three pages with 33 stimulus words on each page. On the top of each page you will find a particular title. This title denotes the task you are supposed to perform. The following are explanations of the tasks:
>
> Superordinates:
> Find a class-name for the stimulus. For instance, class names for FORK are SILVERWARE or UTENSIL. Class names for LIMOUSINE are CAR or VEHICLE.
>
> Following Words:
> Find a word which often follows the stimulus word in a sentence. For example, in the sentence "The pipe belongs to her old father" FATHER, OLD and BELONGS are following words to PIPE.
>
> Parts:
> Name an essential part or attribute of the stimulus. For instance, essential parts of a FORK and the HANDLE, or the METAL. Essential attributes of WISDOM are EXPERIENCE and MATURITY.

Test booklets included in the following order: instructions for free association, free association test form, instructions for restricted associations, restricted association test forms. Order of the stimuli in the test forms was varied so as to minimize the effects of fatigue and of the establishment of patterned associational chains.

Test booklets were distributed to Ss who were instructed to complete them at their leisure, but at one sitting. All booklets were returned within two days after they had been distributed. One week after the first booklets had been returned the second test booklets were given to these same Ss under the same instructions as before. The two booklets were identical except for the order of the restricted association tasks which was counterbalanced. All of the second test booklets were returned within five days after their distribution. Thus, the inter-administration interval varied from seven to fourteen days.

Results

Table 1 indicates, for each task and level of stimulus frequency, the percent of intra-individual reliability over all Ss and stimuli, i.e. the extent to which a S gave the same response on both sessions. The percentages are greatest for Superordinates (49.3%), intermediate for Parts (37.4%), and lowest or Following Words (25.5%). The intra-individual reliability for Free Association (43.1%) lies between that for Superordinates

and Parts. Frequency of the stimulus effected the reliability of the Free, but not of the Restricted Associations. For Free Associations, low frequency words showed higher intra-individual reliability than did high frequency words.

Insert Table 1 about here

As the results of a two-way analysis of variance indicate, the effect of task instructions ($p < .001$) was significant. A series of t-tests between the four tasks indicate that all comparisons were significant except for the difference between the Free Associations and the Superordinates (see first column of Table 2). The difference between the reliability for high and low frequency stimuli under Free Association instruction was not tested for significance since niether the frequency nor the interaction effects were significant. However, the difference appears large enough to suggest further explorations.

Insert Table 2 about here

The question of the degree to which response distributions obtained on the two sessions are alike is more difficult to investigate because there is no single index which adequately allows the comparison of response distributions. This problem of measuring inter-individual reliability is especially great because there are no responses that are more "correct" than others and, thus, traditional summary statistics are difficult to apply and interpret.

As shown in Table 1, the first index by which comparisons were made involves the mean numbers of different responses (types). A three-way analysis of variance indicated a significant difference ($p < .001$) as a function of instructions but not of

frequency or of session. Niether the three-way nor any of the two-way interactions were significant. As shown in Table 1, the mean number of types was lowest (disregarding frequency and session) for the Superordinates (12.8), somewhat higher for Parts (15.1) and highest for Following Words (20.5). Again, the mean for the Free Associations (13.8) was between those for the Superordinates and Parts. Pairwise t-tests indicate significant differences in all cases except between Free Associations and Superordinates, and between Free Associations and Parts (see Table 2). As shown in Table 1 the mean number of common types, i.e. types appearing in the distributions obtained on both sessions, was 7.2 for the Free Associations, 7.3 for Superordinates and Parts, and 7.6 for Following Words. An analysis of variance indicated no significant differences as a function of tasks or frequencies. The number of types and the number of common types are regarded as being negatively correlated with reliability.

Insert Table 3 about here

The nest series of indices were originally developed for comparisons of response distributions of first and second language learners (Ramsye, Riegel and Riegel, 1965). The Venn diagram in Figure 1 depicts graphically the relationships to be discussed. Here, the left-hand circle represents the set of responses obtained in the initial testing and the right-hand circle that of the second testing. The common types occurring an equal number of times on both sessions are represented by MGO (minimum group overlap). The tokens which are common types (responses occurring in both sessions) are denoted by the left and right shaded areas (A + MGO) and (C + MGO) respectively. The tokens which are common types but occur with greater frequency in the first than in the second session are represented by A, while C includes the tokens which are common types and occur more frequently in the second than in the first session. The tokens

which appear in only one distribution are designated B (if they appear in the first session only) and D (if they appear in the second session only). Thus, the higher the reliability, the larger will be the areas MGO, A and C; and the smaller the areas B and D. A serious lack of reliability is easily indicated by increased size of B and D.

Insert Figure 1 about here

An analysis of variance of the minimum group overlap indicated a significant difference ($p < .001$) between tasks but not as a function of stimulus frequency. The interaction was not significant. As shown in Table 3 minimum overlap (MGO), regardless of the frequency of stimuli, is greatest for the Superordinates (68.7%), less for the Parts (60.7%), and least for the Following Words (44.7%). The Free Associations (64.0%) fall between Superordinates and Parts. According to the t-test comparisons of Table 2, all differences between means are significant except for the difference between Free Associations and Superordinates and between Free Associations and Parts.

A three-way analysis of variance on the response sets (A + MGO) and (C + MGO) revealed a significant difference between the tasks ($p < .01$); but neither frequency, sessions, the three-way interaction, nor any of the two-way interactions were significant. As can be derived from Table 3, the total percentages of overlaping responses (A + MGO and C + MGO) are highest for the Superordinates (79.0 and 79.7), slightly lower for Parts (72.7 and 71.3), and lowest for the Following Words (53.3 and 54.3). As before, Free Associations (74.3 and 77.3) fall between those of the Superordinates and Parts. Throughout, the differences between the response sets (A + MGO) and (C + MGO) are negligible. As the t-test comparisons of Table 2 indicate all differences are significant ($p < .01$) except those between Free Associations and Superordinates

and between Free Association and Parts.

A last comparison between tasks involves the proportion of (A + MGO) and (C + MGO) accounted for by MGO. For Free Association this proportion is 84.0%, for Superordinates 86.0%, for Parts 83.3%, and for Following Words 83.0%, thus indicating no task differences despite the fact that A and C, and MGO varied significantly between tasks.

Discussion

Although this study is concerned with reliability, traditional measures, such as test-retest correlations are inappropriate because the present data emphasize response distributions and do not fit into the dichotomous categories of "correct" and "incorrect" answers. Subsequently, simple enumerations of differences in the response distributions and of the identical responses given at the two testing sessions have been used for the evaluations.

As predicted, reliability differed as a function of the instructions. Superordinates were most reliable, Parts somewhat less, and Following Words least reliable. The differences were statistically significant in all instances. Though no specific prediction was made, the reliability of Free Associations fell consistently between those for the Superordinates and Parts, in some cases differing significantly from one or the other. Using the findings of this study (see Table 3) the Venn diagram in Figure 1 may be redrawn to more accurately reflect the proportions of overlap.

Possibly, the differences in reliability can be accounted for by two negatively correlated factors--mean strength of primary response and mean number of different responses (types). In an attempt to investigate some of the relationships between the different measures of reliability, the primary response strengths and the number of types as determined from the Mighigan Restricted Association Norms (Riegel, 1965) have been correlated, separately for each task, with the percentage of tokens which were

common types (area A + MGO and C + MGO of Figure 1 and Table 3).

For the primary response strengths the correlations were .63 for Superordinates, .76 for Parts and .57 for Following Words. For Free Association the correlation was only .23 thus indicating that, even though reliability is relatively high, strength of the primary response is not nearly as great a determinant of reliability for this condition as compared to the restricted conditions. For the number of types the correlations were -.82 for Superordinates, -.79 for Parts, and -.85 for Following Words, indicating that for each of the restricted tasks the fewer the number of types the higher the reliability. For Free Association the correlation was -.58 suggesting again that the number of types is not as important a determinant of reliability in Free Association as it is in Restricted Associations. Table 4 gives the mean strength of the primary, secondary and tertiary responses.

Insert Table 4 about here

Although no significant differences in reliability were found as a function of stimulus frequency, the slight differences suggest further research on this matter. Special consideration should be given to the following factors: First, the frequency classification made in this study was very gross. If frequency is in any sense a factor to be reckoned with, finer distinctions will have to be investigated. Second, the observed differences are congruent with the results of Riegel and Riegel (1964) which suggest a curvilinear relationship between word frequency and primary response strength (and hence with reliability which appears, at least in Restricted Associations, to be correlated with primary response strength).

As to the absolute magnitude of the reliability of the Norms, the intra-individual index and the index of tokens which are common types are the most illuminating measures

for they are proportions based upon consideration of all the responses in the distributions. Although there is no way of arbitrarily defining a cut-off point of reliability, the scores on these indices are sufficiently high to regard the Restricted Associations as reliable. An experimenter who intends to use the Michigan Restricted Associations ought to take the differences between the tasks into account. His decision will, of course, be also based on the particular demands which his study implies. The reliability estimates reported here are expected to prove useful as a reference point for future comparisons of the overlaps between the response distributions of groups differing with respect to their languages, language proficiency, educational and social status, pathology, etc. Further research is in progress on long term reliability and reliability as a function of sex, form class of the stimuli, task category (logical, intra-logical, grammatical), and other specific tasks.

Summary

In this study the reliability of the Michigan Restricted Association Norms was estimated. Thirty Ss were asked to respond to 33 stimuli under each of four instructions and to perform the identical tasks seven to fourteen days later. Comparisons of the intra-individual reliability and of the response distributions (inter-individual reliability) were made between tasks, frequency of stimuli, and sessions. The distributions differed markedly between the tasks but not between stimulus frequencies and sessions. The absolute magnitude of the reliability can be regarded as sufficiently high though such an evaluation will have to depend on the particular purpose for which the Norms are to be employed. Reliability is highest for Superordinates, slightly lower for Free Associations and Parts, and lowest for Following Words. Attempts were also made to compare the differences in reliability with differences in the response distributions of the Norms. or Restricted but not for Free Associations the strength of the primary responses in the Michigan Norms correlates with the reliability coefficients.

References

DiVesta, F. J. Norms for modifiers used by children in a restricted word association task. Report No. 9, USPHS Grant HD-00872, Pennsylvania State University, 1964.

Gekoski, W. L. The adequacy of free association norms, umpublished manuscript, 1964.

Jenkins, J. J. and Palermo, D. S. Contrast responses to the stimulus words of the Palermo-Jenkins word association list. Res. Bull. No. 43, USPHS Grant HD-00961, Pennsylvania State Univ., 1964a.

Jenkins, J. J. and Palermo, D. S. Subordinate responses to the stimulus words of the Palermo-Jenkins word association lists. Res. Bull. No. 46, USPHS Grant HD-00961, Pennsylvanis State Univ., 1964b.

Jenkins, J. J. and Palermo, D. S. Further changes in word association norms. J. person. soc. Psychol., 1965, 1, 303-309.

Jenkins, J. J. and Russell, W. A. Systematic changes in word association norms, 1910-1952. J. abnorm. soc. Psychol., 1960, 60, 293-304.

Jung, C. G. Studies in Word Associations. London: Heinemann, 1918.

Kent, G. H. and Rosanoff, A. J. A study of association in insanity. Amer. J. Insanity, 1910, 67, 37-96, 317-390.

Martin, D. R. and Clifton, C. A normative collection of opposite responses. Rep. No. 14, NSF Grant G-18690, Univ. Minnesota, 1964.

May, M. A. The mechanism of controlled association. Arch. Psychol., 1917, 5, 1-74.

Ramsey, R. M., Riegel, K. F., and Riegel, Ruth M. A comparison of first and second language learning of American and Spanish students. Rep. No. 4, NIMH Grant No. MH 07619 - OIAI, Univ. Michigan, 1965.

Riegel, K. F. The Michigan restricted association norms. Rep. No. 3, NIMH Grant No. MH 07619 - OIAI, Univ. Michigan, 1965.

Riegel, K. F. Age and cultural differences as determinants of word associations: suggestions for their analysis. Psychol. Rep., 1965, 16, 75-78.

Riegel, K. F. and Riegel, Ruth M. Changes in associative behavior during later years of life: a cross-sectional analysis. Vita Humana, 1964, 7, 1-32.

Rosanoff, I. R. and Rosanoff, A. J. A study of association in children. Psychol. Rev., 1913, 20, 43-89.

Thorndike, E. L. and Lorge, J. The Teacher's Word Book of 30,000 Words. New York: Bureau of Publ., Teachers College, Columbia University, 1944.

Watt, H. J. Experimentelle Beiträge zu einer Theorie des Denkens. Arch. ges. Psychol., 1905, 4, 289-436.

Woodrow, H. and Lowell, F. Children's association frequency tables. Psychol. Monogr. 1916, 22, Whole No. 97.

Wreschner, A. Die Reproduktion von Association und Vorstellungen. Z. Psychol. Ergbl., 1907, No. 3.

Table 1

Mean percent intra-individual reliability and mean number of types and common types for four tasks, two frequency levels, and two sessions.

		Intra-individual reliability	Types Session 1	Types Session 2	Common Types
Free Association	High	37.6	14.9	14.2	7.4
	Low	49.8	13.2	12.5	7.0
Superordinates	High	49.4	12.9	12.6	7.3
	Low	49.1	12 9	12.6	7.3
Parts	High	38.1	15.2	15.2	7.3
	Low	36.4	15.2	14.7	7.3
Following Words	High	24.3	21.4	20.7	7.4
	Low	26.9	20.4	19.5	7.7

Table 2

Significance levels of differences between means on four indices of reliability

	Intra-individual reliability	Types**	Minimum Overlap*	Group Overlap*
Free Association-Superordinates	NS	NS	NS	NS
Free Association-Parts	.05	NS	NS	NS
Free Association-Following Words	.01	.01	.01	.01
Superordinates-Parts	.01	.01	.025	.01
Superordinates-Following Words	.01	.01	.025	.01
Parts-Following Words	.01	.01	.01	.01

* df = 64 (As shown in Figure 1, Minimum Overlap = MGO; Group Overlap = Average of A + MGO and C + MGO.)

** df = 132

Table 3

Indices of response overlaps in percentages

as explained in Figure 1.

	MGO	A	B	C	D
Free Association	54.0	10.3	25.7	13.3	22.7
Superordinates	68.7	11.7	21.0	11.0	20.3
Parts	60.7	12.0	27.3	10.6	28.7
Following Words	44.7	8.6	46.7	9.6	45.7

Table 4

Mean strength of the three highest ranking responses and mean number of types determined from the Michigan Restricted Association Norms for the stimuli in the present study.

	Primary	Secondary	Tertiary	Types
Free Association	31.2	13.2	7.8	33.3
Superordinates	36.3	12.4	7.5	28.7
Parts	27.8	12.8	7.9	36.7
Following Words	15.7	9.0	5.5	54.6

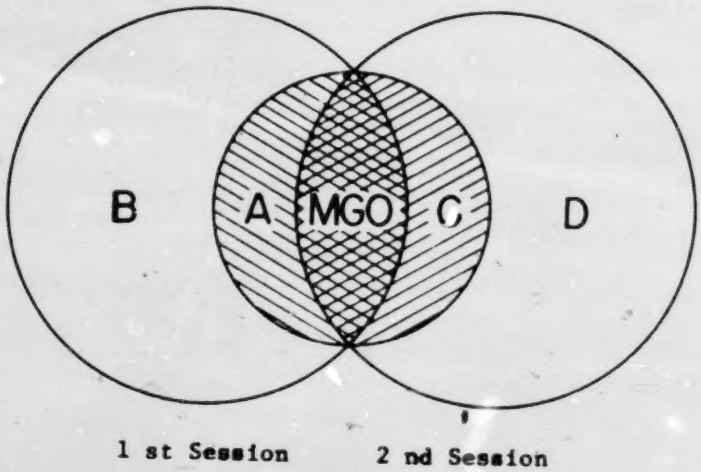

Fig. 1. Subsets of responses given on the two sessions. (MGO = common types occurring an equal number of times in both sessions: A + MGO = tokens on the first session which are common types; C + MGO = tokens on the second session which are common types; B = tokens which are types that occur on the first but not the second session; D = tokens which are types that occur on the second but not the first session. Both A + B + MGO and C + D + MGO equal 100 percent.)